



# Full motion tracking in ultrasound using image speckle information and visual servoing

A. Krupa, G. Fichtinger, G. Hager

## ► To cite this version:

A. Krupa, G. Fichtinger, G. Hager. Full motion tracking in ultrasound using image speckle information and visual servoing. IEEE Int. Conf. on Robotics and Automation, ICRA'07, 2007, Roma, Italy. pp.2458-2464. inria-00350650

**HAL Id: inria-00350650**

**<https://hal.inria.fr/inria-00350650>**

Submitted on 7 Jan 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Full Motion Tracking in Ultrasound Using Image Speckle Information and Visual Servoing

Alexandre Krupa, Gabor Fichtinger and Gregory D. Hager

**Abstract**—This paper presents a new visual servoing method that is able to stabilize a moving area of soft tissue within an ultrasound B-mode imaging plane. The approach consists of moving the probe in order to minimize the relative position between a target imaging plane and the ultrasound plane observed by the probe of the moving tissue target. The problem is decoupled into motion out-of-plane and motion within plane. For the former, a new original method based on the speckle information contained in the images is developed. For the latter, an image region tracker is used to provide the in-plane motion. A visual servoing control scheme is then developed to perform the tracking robotic task. The method is validated on simulated motions of a probe on a static ultrasound volume acquired from a phantom.

**Index Terms**—Visual servoing, ultrasound, speckle correlation, medical robotics.

## I. INTRODUCTION

Over the last several years, there have been a variety of investigations of methods to assist ultrasound (US) guided interventions or ultrasound examination by the use of medical robotic systems. However, only a reported systems use the visual information provided by the ultrasound sensor directly in their control scheme. In [1], visual servoing is used to center a section of an artery within the 2D US image (B-scan image) by controlling the 3 degrees of freedom (DOF) in the observation plane (2 translations and 1 rotation). In [2], the authors present a robotic system including a motionless ultrasound probe and a 2-DOF needle manipulator. Automatic needle insertion in a soft sponge phantom was performed using ultrasound image-based visual servoing. However, in this work, the actuated needle must lie in the ultrasound observation plane since only 2 DOF inside the observation plane are controlled. In general, a standard probe provides a 2D B-scan image which therefore limits vision-based control to the 3 DOF contained in the plane using classic visual servoing techniques. Of course, this is not the case if a 3D ultrasound probe is used, as in [3] where a surgical instrument is positioned by 3D visual servoing. However, for the moment 3D ultrasound probes are expensive and provide only low voxel resolution. In what follows we will only consider the use of standard 2D probes which are more widespread in medical centers.

Some recent studies eliminate the requirement of controlling only the DOF contained in the ultrasound plane. In [4]

4 DOF, which are not necessary in the observation plane, are controlled by visual servoing in order to automatically position a robotized laparoscopic instrument. In [5] a visual servoing technique was developed to reach a desired section of a tumor by controlling the 6-DOF motion of the probe. However, these methods need a geometrical model of the object of interest to interact with, which is the tool forceps in [4] and a pre-operative model of the tumor in [5]. These methods also need extensive and complex image processing to segment the targeted objects in the B-scan images.

In this paper we propose to perform 6 DOF control of a medical robot by using visual servoing from ultrasound B-scan images. Rather than relying on segmenting objects of interest, we make direct use of the speckle information contained in the image. Using speckle, we track both out-of plane and in-plane motions. Usually speckle contained in ultrasound images is considered to be a noise and much effort has been devoted to reducing it for image segmentation applications. However speckle is not a random noise, such as white noise, but instead is highly correlated over small motions of the probe. This follows from the fact that speckle results from the coherent reflection of very small cells contained in soft tissues. Moreover, in practice focusing of the ultrasound beam is not perfect, and particularly for the elevation direction (that is, the direction orthogonal to the imaging plane). Since this induces a ultrasound plane width of a couple of mm there is an overlap between cells observed in close B-scans. The result is a speckle correlation between successive ultrasound images captured along the elevation direction. Several authors ([6], [7]) have published speckle decorrelation techniques for performing freehand 3D ultrasound imaging without the need of a position sensor to provide the location of the ultrasound probe. In [8], a method using speckle information was also proposed to control on-line the quality of the ultrasound calibration needed for freehand 3D imaging. It is an ultrasound calibration technique based on a closed formulation on which the in-plane motion of the ultrasound plane was estimated with speckle correlation. In [9], strain imaging of tumors was performed thanks to correlation based on speckle.

As a result of our methods, we show that it is possible to automatically track a moving 3D ultrasound volume by an ultrasound probe actuated by a robot. This can be useful in numerous medical applications, for example to automatically move the probe to maintain an appropriate view of moving soft tissues during an examination or to synchronize the insertion of a needle with a moving tumor for biopsy or therapy purpose. In this paper, the robotic task consists of

Alexandre Krupa is with IRISA - INRIA Rennes, Campus de Beaulieu, 35042 Rennes Cedex, France alexandre.krupa@irisa.fr and has been in sabbatical in Johns Hopkins University

Gabor Fichtinger and Gregory D. Hager are with the Engineering Research Center, Johns Hopkins University, Baltimore, MD 21218, USA {gabor,hager}@cs.jhu.edu

automatically moving the current B-scan image provided by the probe in order to track a moving target B-scan previously recorded by the user at a given location. In our approach the 3-DOF out-of-plane and 3-DOF in-plane motions are respectively extracted by an estimation method using the speckle information and an image region tracking algorithm based on grey level intensity. Two independent visual servoing based control schemes on the extracted motion are then proposed to control all DOF of the probe.

The remainder of this paper is organized as follows: Sections II and III present respectively the methods used to extract the out-of-plane motion and the in-plane motion of the target B-scan image. The visual servoing control laws are developed in section IV and results obtained from simulations are presented and discussed in section V.

## II. OUT-OF-PLANE MOTION EXTRACTION

### A. Speckle decorrelation technique

In our approach, we propose to estimate the out-of-plane motion of an ultrasound image plane target with respect to the current B-scan image by using image speckle information. Out-of-plane motion includes the displacement along the normal to the current B-scan image and the two rotations around its axial and lateral axes.

The source of motion information in these directions will be speckle decorrelation. An approximation of the speckle correlation function as a function of the orthogonal distance  $d$  between two B-mode scans  $\mathbf{I}_1$  and  $\mathbf{I}_2$  is given in [6] by the Gaussian model function:

$$\rho(\mathbf{I}_1, \mathbf{I}_2) = \exp\left(-\frac{d^2}{2\sigma^2}\right) \quad (1)$$

where  $\rho$  is the correlation value of speckle included in two corresponding patches in the two images and  $\sigma$  is the resolution cell width along the elevation direction. In practice, this approximation works well when the grey level intensity of the image is defined on a linear scale. This is the case when we directly use the radio-frequency (RF) signal provided by the ultrasound imaging device. Unfortunately, this signal is not generally available on most standard ultrasound systems. Instead, the RF data is processed into B-mode images with intensity compressed on a logarithmic scale. As this is our case, we first have to convert the intensity to a linear scale by applying the relation given in [10]:

$$I(i, j) = 10^{P(i, j)/51} \quad (2)$$

where  $I(i, j)$  is the decompressed grey level intensity of the pixel located at image coordinates  $i, j$  and  $P(i, j)$  is the measured intensity in the B-mode image.

In order to perform position estimation using decorrelation, it is necessary to experimentally calibrate speckle decorrelation curves from real soft tissues or from an ultrasound phantom simulating speckle. These curves are obtained by capturing a set of B-scan images at known distances along

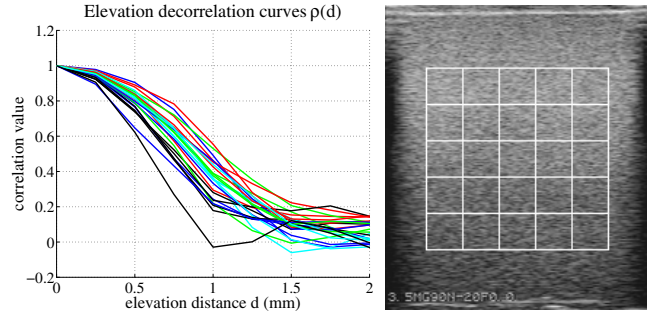


Fig. 1. (left) Experimental decorrelation curves of the 25 patches considered in the (right) ultrasound image

the elevation direction and measuring the normalized correlation coefficients given by  $\rho(d) =$

$$\frac{\sum_{i=1}^m \sum_{j=1}^n (I_0(i, j) - \bar{I}_0)(I_d(i, j) - \bar{I}_d)}{\sqrt{\sum_{i=1}^m \sum_{j=1}^n (I_0(i, j) - \bar{I}_0)^2 \sum_{i=1}^m \sum_{j=1}^n (I_d(i, j) - \bar{I}_d)^2}} \quad (3)$$

for several patches positioned in the images. In (3),  $I_0$ ,  $I_d$  correspond respectively to the pixel intensity array of a given patch of the B-scan image captured at  $d = 0$  and the one of the corresponding patch in the image captured at distance  $d$ .  $\bar{I}_0$ ,  $\bar{I}_d$  are the mean value intensity of these patches and  $m$  and  $n$  are their height and width. Fig. 1 shows the decorrelation curves when we consider a grid of 25 patches in the image.

As described with (1), the observed decorrelation curves behave like Gaussian functions, but with different parameters  $\sigma$ . This is due to the fact that the resolution cell width  $\sigma$  is a function of the lateral and axial position of the patch in the image. Generally for sensorless freehand 3D ultrasound, a look-up table based on these calibrated decorrelation curves is used to provide an accurate estimation of the elevation distance from the considered measured inter-patch correlation value. In our tracking application the objective is to minimize the relative position between the current B-scan and a desired one, therefore we do not require high accuracy on the target plane position estimation. Consequently we propose to estimate the inter-patch elevation distance directly from (1) by using:

$$\hat{d}(\rho) = \sqrt{-2\hat{\sigma}^2 \ln(\rho)} \quad (4)$$

where  $\hat{\sigma} = 0.72$  mm is identified by averaging the experimental decorrelation curves and fitting the model function.

### B. Plane estimation

To estimate the target plane position, a minimum of three elevation distances between three non-collinear patches are needed. As (4) gives only absolute value on  $d$ , we have to determine the correct sign of each distance. If we first assume that the sign of each inter-patch distance is known, we can estimate the target plane position with respect to the current B-scan by using the plane equation:

$$ax + by + cz + d = 0 \quad (5)$$

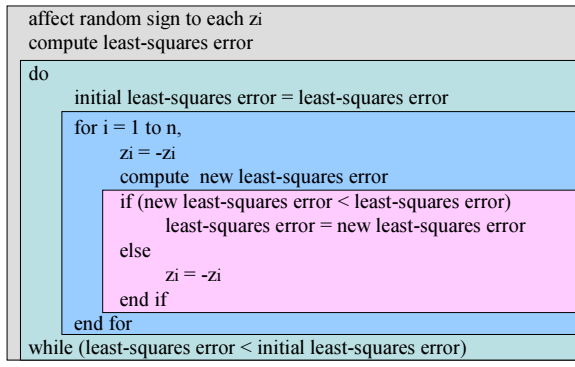


Fig. 2. Iterative algorithm for plane position estimation

where  $x, y, z$  are the 3D coordinates of the center of a patch belonging to the target image plane with respect to the current image plane.  $x, y$  correspond to its 2D position fixed in the image grid (the same for the current and target image plane) and  $z$  is the signed elevation distance which can be estimated from (5) by:

$$\hat{z} = \sum_{j=1}^3 \alpha_j f_j(x, y) \quad (6)$$

where  $f_1(x, y) = 1$ ,  $f_2(x, y) = x$ ,  $f_3(x, y) = y$  depend on the coordinates  $x, y$  which are known and  $\alpha_1 = -d/c$ ,  $\alpha_2 = -a/c$ ,  $\alpha_3 = -b/c$  are the parameters of the plane. By considering all the  $n$  patches of the grid, these parameters can be estimated by using a classical least-squares algorithm whose the cost function to minimize is the sum of squares of the differences between the estimated and observed elevation distances:

$$J = \sum_{i=1}^n (\hat{z}_i - z_i)^2 \quad (7)$$

and which gives solution:

$$(\alpha_1 \ \alpha_2 \ \alpha_3)^T = (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \mathbf{Z} \quad (8)$$

where the components of the  $n \times 3$  matrix  $\mathbf{M}$  are given by  $M_{i,j} = f_j(x_i, y_i)$  with  $i = 1 \dots n$ ,  $j = 1 \dots 3$  and the vector  $\mathbf{Z}$  contains the  $n$  observed elevation distances  $Z_i = z_i$ . The normal vector of the target plane expressed in the current plane is then obtained by:

$$\vec{n} = (a \ b \ c)^T = \frac{(\alpha_2 \ \alpha_3 \ 1)^T}{\|(\alpha_2 \ \alpha_3 \ 1)^T\|} \quad (9)$$

and the elevation distance between the centers of the current and target B-scan images is  $z_b = \alpha_1$ .

This least-squares algorithm cannot be directly applied to estimate the plane position due to the sign ambiguity of the  $z_i$  distance of each patch. So we propose to use the iterative algorithm presented in Fig. 2 to rearrange sign of each distance. The principle is first choose a random sign on each  $z_i$  and to compute an initial plane estimate and least-squares error using these signs. Then, we modify the sign of a patch and compute the new least-squares error. If the new error norm is lower than the previous one, then the sign

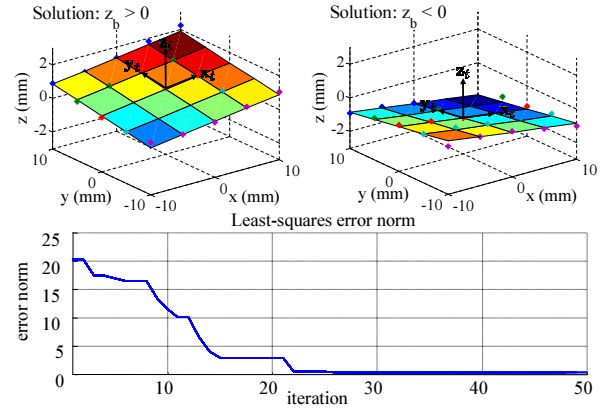


Fig. 3. (top) Symmetric plane position solutions provided by the iterative algorithm. The points on the planes show the rearranged (signed) positions of patches after the algorithm convergence. (bottom) Plots of the decreasing least-squares error norm during the iterative algorithm process

is kept or otherwise it is discarded. This process is repeated for the  $n$  patches in a loop. At the end, if the resulting error norm is lower than the initial one then the initial error is set to the current one and the loop is repeated until the last resulting error is the same as the initial error. The algorithm will then stop when it converges to one of the two stable symmetric solutions as illustrated in Fig. 3. The first solution corresponds to the case when there is a positive elevation distance  $z_b > 0$  between the target and current plane and the second to the case for a negative distance  $z_b < 0$ . Note that from one solution we can easily determine the second. For the case presented in Fig. 3, the algorithm converges only with 50 iterations whereas there are, in principle,  $2^n$  (with  $n = 25$ ) possible configurations<sup>1</sup> of the signed distances.

Now, if we consider a 3D Cartesian frame  $\{c\}$  attached to the center of the current image plane and another  $\{t\}$  attached to the center of the target image plane, the relative displacement due to the out-of-plane motion can be defined as a translation of distance  $z_b$  along the  $Z$  axis of  $\{c\}$  and two successive rotations  $\alpha$  and  $\beta$  around the  $Y$  and  $X$  axes. This gives us the following homogeneous transformation matrix between  $\{c\}$  and  $\{t\}$ :

$${}^c\mathbf{H}_t = \begin{pmatrix} \cos(\alpha) & \cos(\alpha)\sin(\theta) & \sin(\alpha)\cos(\theta) & 0 \\ 0 & \cos(\theta) & -\sin(\theta) & 0 \\ -\sin(\alpha) & \cos(\alpha)\sin(\theta) & \cos(\alpha)\cos(\theta) & z_b \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (10)$$

As the third column corresponds to the  $Z$  axis of the target plane expressed in the current frame  $\{c\}$  the angles  $\alpha$  and  $\beta$  can be directly determined from the components of the estimated normal vector  $\vec{n}$ . Therefore the 2 solutions of  ${}^c\mathbf{H}_t$  are given by using:

$$\begin{aligned} \alpha &= \text{atan}(a/c), \theta = -\text{asin}(b) \text{ if } z_b > 0 \\ \alpha &= \text{atan}(-a/c), \theta = -\text{asin}(-b) \text{ if } z_b < 0 \end{aligned} \quad (11)$$

<sup>1</sup>In fact, there are fewer than  $2^n$  due to the planarity constraint; indeed this is why such a simple algorithm works.

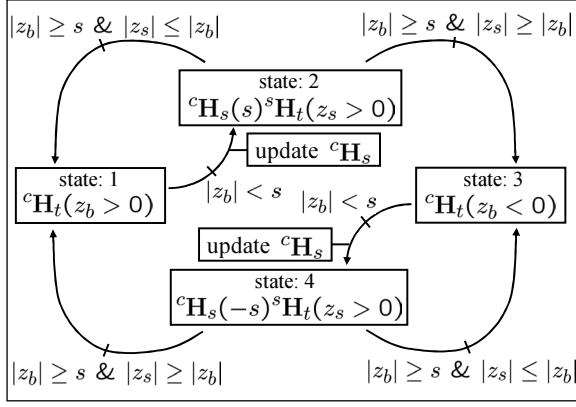


Fig. 4. The state-transition graph used to track the sign of the elevation distance  $z_b$  and compute the relative position  ${}^c\mathbf{H}_t$  between the current and target planes

Note that, if  $z_b = 0$  there is an ambiguity on the target plane orientation. This problem will be considered next.

### C. Elevation distance tracking

Once a correct sign is known for the elevation plane, it is possible to develop a system for tracking it without the need for continual estimation. This approach also provides a means for extending the range of the elevation distances that can be corrected for.

In order to resolve the remaining sign ambiguity and initiate tracking, we have developed a state-transition graph which memorizes the evolution of the sign and uses an intermediate B-scan image to reconstruct the target frame position  ${}^c\mathbf{H}_t$  when  $|z_b|$  is close to zero. In practice, the B-scan image target that needs to be tracked will be initially set from a desired ultrasound image. This will be done after an initial procedure where the user positions the probe held by a medical robot to see the image of interest. Therefore at the start, the current image and the target B-scan are superposed, so  $z_b = 0$ . We then propose to initially move the probe by a small control step in the negative elevation direction in order to obtain  $z_b > s$  where  $s$  is a very low threshold value. This provides initialization for the state-transition graph presented in Fig. 4.

In particular, this first motion provides data for state 1 where the position of the target is given by  ${}^c\mathbf{H}_t(z_b > 0)$ . This state is maintained while  $z_b > s$ . If  $|z_b|$  decreases below the threshold  $s$  due to the motion of soft tissues then an intermediate plane with Cartesian frame  $\{s\}$  is set and frozen to the current target B-scan position  ${}^c\mathbf{H}_s = {}^c\mathbf{H}_t(s)$  and the state switches from 1 to 2. In this new state the position of the plane target is then given by  ${}^c\mathbf{H}_t = {}^c\mathbf{H}_s(s)^s\mathbf{H}_t(z_s > 0)$  where  ${}^s\mathbf{H}_t(z_s > 0)$  is the homogeneous matrix from the fixed intermediate plane to the target plane provided by the estimation method described in section II-B and  $z_s$  is the corresponding elevation distance between these two planes.

This new state is maintained while  $|z_b| < s$ . Of course there is the possibility of going back to the state 1 if  $z_b$  increases when the transition  $|z_b| \geq s$  &  $|z_s| \leq |z_b|$  is

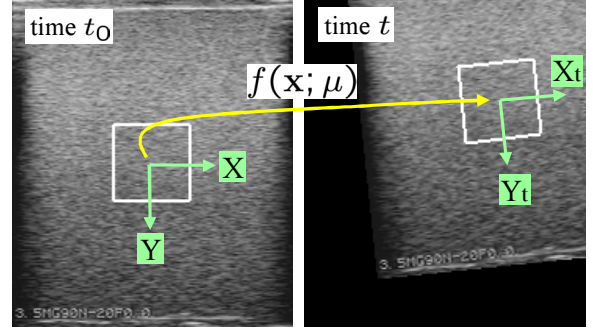


Fig. 5. (left) Reference image acquired at time  $t_0 = 0$  with the region of interest to track - (right) current image modified by the in-plane motion  $f(\mathbf{x}; \mu)$  with the estimated region of interest

validated. If now  $|z_b| \geq s$  &  $|z_s| \geq |z_b|$  which means that  $z_b$  is negative and is lower than  $-s$  then the state goes in 3 where the target position is given directly by the solution with negative elevation distance  ${}^c\mathbf{H}_t(z_b < 0)$ . If afterwards  $|z_b|$  becomes lower than the threshold, the intermediate plane is updated and frozen to the current target position  ${}^c\mathbf{H}_s = {}^c\mathbf{H}_t(-s)$  and the state goes to 4 with solution  ${}^c\mathbf{H}_t = {}^c\mathbf{H}_s(-s)^s\mathbf{H}_t(z_s > 0)$  where  ${}^s\mathbf{H}_t(z_s > 0)$  is the transformation matrix from the recent updated intermediate plane to the target. The first state is then retrieved when  $|z_b| \geq s$  &  $|z_s| \geq |z_b|$ . This method permits computation of the correct sign of the distance  $z_b$  by taking into account its evolution and avoiding ambiguous orientation case when  $z_b = 0$ . Moreover in order to obtain smooth transitions when the state switches, the following interpolation function is applied to give the target plane pose vector  $\mathbf{p}$ :

$$\mathbf{p} = (1 - (|z_b|/s)^2) \mathbf{p}_1 + (|z_b|/s)^2 \mathbf{p}_2 \quad (12)$$

where  $\mathbf{p}_1$  is the pose vector describing the reconstructed homogeneous matrix  ${}^c\mathbf{H}_t$  obtained during state 2 or 4 and  $\mathbf{p}_2$  is the pose vector describing the direct solution  ${}^c\mathbf{H}_t$  during state 1 or 3. Note that this function gives no weight to the direct solution  ${}^c\mathbf{H}_t$  when  $z_b = 0$  in order to reject the unstable case. The components of the normal vector  $\vec{n}$  of the B-scan plane and its orientation angles  $\alpha, \beta$  are then retrieved using (10) and (11).

### III. IN-PLANE-MOTION-EXTRACTION

Fig. 5 shows the target image captured at time  $t_0 = 0$  and the current image obtained at time  $t$  when in-plane motion was applied. To extract the in-plane rigid motion of the B-scan target with respect to the current image plane, we use the image region tracking technique presented in [11] which we briefly recall here.

The objective of this technique is to estimate the parameter vector  $\mu$  of an appropriate parametric model function  $f(\mathbf{x}; \mu)$  which describes the geometrical transformation on the pixel coordinates  $\mathbf{x} = (x \ y)^T$  from the reference to the current image. For our application, the motion parameter vector is  $\mu = (u_x \ u_y \ \gamma)^T$  where  $u_x, u_y$  are the translations along  $X$  and  $Y$  axes of the reference image and  $\gamma$  is the rotation angle

around the  $Z$  axis. The vector form of the motion parametric model function is:

$$f(\mathbf{x}; u_x, u_y, \gamma) = \mathbf{R}(\gamma)\mathbf{x} + \mathbf{u} \quad (13)$$

where  $\mathbf{R}(\gamma)$  is the  $2 \times 2$  rotation matrix of angle  $\gamma$  and  $\mathbf{u} = (u_x \ u_y)^T$  is the translational vector. The principle of the method is to determine the motion parameter  $\mu$  which minimizes the sum of squared differences of pixels intensity between the region of interest obtained with the geometrical transformation (13) in the current image and the reference region of interest fixed in the target image where  $\mu = 0$ . Therefore the objective function to minimize is:

$$\mathcal{O}(\mu) = \|\mathbf{I}(\mu, t) - \mathbf{I}(0, t_0)\|^2 \quad (14)$$

where  $\mathbf{I}(0, t_0)$  is the vector containing the intensity values of the  $N$  pixels belonging to the reference target image at  $t = 0$  and  $\mathbf{I}(\mu, t)$  contains the intensity values of the  $N$  pixels in the current image whose coordinates were warped by (13) using the most recent motion parameter  $\mu(t)$  as given here:

$$\mathbf{I}(\mu, t) = \begin{bmatrix} I(f(\mathbf{x}_1, \mu), t) \\ \vdots \\ I(f(\mathbf{x}_N, \mu), t) \end{bmatrix} \quad (15)$$

By rewriting (14) in term of a vector of offsets  $\delta\mu$  such that  $\mu(t + \tau) = \mu(t) + \delta\mu$  from an image captured at time  $t + \tau$ :

$$\mathcal{O}(\delta\mu) = \|\mathbf{I}(\mu + \delta\mu, t + \tau) - \mathbf{I}(0, t_0)\|^2 \quad (16)$$

and by approximating it with a first order Taylor expansion we obtain:

$$\mathcal{O}(\delta\mu) \approx \|\mathbf{M}\delta\mu + \mathbf{I}(\mu, t + \tau) - \mathbf{I}(0, t_0)\|^2 \quad (17)$$

where  $\mathbf{M}$  is the Jacobian matrix of  $\mathbf{I}$  with respect to  $\mu$ :

$$\mathbf{M}(\mu) = \begin{bmatrix} \nabla_{\mathbf{x}} I(\mathbf{x}_1, t_0)^T f_{\mathbf{x}}(\mathbf{x}_1, \mu)^{-1} f_{\mu}(\mathbf{x}_1, \mu) \\ \vdots \\ \nabla_{\mathbf{x}} I(\mathbf{x}_N, t_0)^T f_{\mathbf{x}}(\mathbf{x}_N, \mu)^{-1} f_{\mu}(\mathbf{x}_N, \mu) \end{bmatrix} \quad (18)$$

Here  $\nabla_{\mathbf{x}} I(\mathbf{x}, t_0)^T$  is the intensity gradient vector at pixel location  $\mathbf{x} = (x \ y)^T$  in the reference image and  $f_{\mathbf{x}}, f_{\mu}$  are respectively the partial derivatives of  $f(\mathbf{x}; \mu)$  with respect to  $\mathbf{x}$  and  $\mu$ . By using  $\mu = (u_x \ u_y \ \gamma)^T$  and the parametric motion model (13) it gives:

$$f_{\mathbf{x}}^{-1} f_{\mu} = \begin{bmatrix} 1 & 0 & -y \\ 0 & 1 & x \end{bmatrix} \begin{bmatrix} \mathbf{R}(-\gamma) & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \quad (19)$$

The solution of  $\delta\mu$  is then obtained by setting the gradient of  $\mathcal{O}(\delta\mu)$  to zero and solving which yields:

$$\delta\mu = -\mathbf{M}^+(\mathbf{I}(\mu, t + \tau) - \mathbf{I}(0, t_0)) \quad (20)$$

where  $\mathbf{M}^+$  is the pseudo inverse of  $\mathbf{M}$ . The motion parameter vector is then:

$$\mu(t + \tau) = \mu(t) + \delta\mu \quad (21)$$

In practice, in order to obtain accurate parameters convergence, we successively compute (20) and (21) during several iterations until  $\|\delta\mu\|^2$  becomes lower than a small

fixed threshold value  $\epsilon$ . For more complete details on this method we invite the reader to refer to [11]. The extracted motion is then applied to the  $n$  patches used to compute the speckle correlation in order to update their positions in the current image. Other methods based on the same principle are proposed in the literature, for example in [12] a second-order minimization technique is used for large motion tracking with fast convergence rate by using the mean value of the Jacobian  $\mathbf{M}$  in the target image and the one in the current image.

#### IV. VISUAL SERVOING

Now that the whole motion of the B-scan target is estimated, we present the control scheme used to control a medical robot holding the ultrasound probe in order to reach and track a moving B-scan target. In our approach, two visual servoing techniques are used to independently control the out-of-plane 3-DOF motion of the probe and its in-plane 3-DOF motion.

##### A. Out-of-plane motion control

A 3D visual servoing control is used to track the out-of-plane motion. We chose as the visual features  $\mathbf{s}_1 = (a \ b \ c \ z_b)^T$  the 3 components of the normal vector  $\vec{n}$  of the estimated target plane and its elevation distance  $z_b$  with respect to the current B-scan. The desired visual features vector to achieve is  $\mathbf{s}_1^* = (0 \ 0 \ 1 \ 0)^T$  which means that the final position of the normal vector of the target plane will be orthogonal to the current image and that relative elevation distance will be null. The variation of the visual information  $\mathbf{s}_1$  to the out-of-plane velocity  $\mathbf{v}_1 = (v_z \ \omega_x \ \omega_y)^T$  of the probe is given by:

$$\dot{\mathbf{s}}_1 = \mathbf{L}_{s_1} \mathbf{v}_1 = \mathbf{L}_{s_1} \begin{bmatrix} 0 & 0 & -c \\ 0 & c & 0 \\ 0 & -b & a \\ -1 & 0 & 0 \end{bmatrix} \mathbf{v}_1 \quad (22)$$

where  $v_z$  is the probe translational velocity along the orthogonal  $Z$  axes of the current image frame  $\{c\}$  (attached to the center of the image) and  $\omega_x, \omega_y$  are respectively the rotational velocities around the  $X$  and  $Y$  axis. In visual servoing  $\mathbf{L}_{s_1}$  is called the interaction matrix (see [13]) and is determined from the geometrical model of the considered system. In our case it depends only on the components of the normal vector  $\vec{n}$  of the target plane. The visual servoing task can then be expressed as a regulation to zero of the task function  $\mathbf{e}_1 = \mathbf{s}_1 - \mathbf{s}_1^*$ . Usually, the control law is defined such as the task  $\mathbf{e}_1$  decreases exponentially in order to behave like a first order system by using a proportional controller [13]. In this work we apply rather the second-order minimization technique introduced in [14] which uses the following control law to improve the trajectory for large displacement:

$$\mathbf{v}_1 = -2\lambda_1(\widehat{\mathbf{L}}_{s_1} + \mathbf{L}_{s_1}^*)^+ \mathbf{e}_1 \text{ with gain } \lambda_1 > 0 \quad (23)$$

where  $\widehat{\mathbf{L}}_{s_1}$  is the interaction matrix estimated at each control iteration and  $\mathbf{L}_{s_1}^*$  is the one at the desired location (with  $a = b = 0$  and  $c = 1$ ).



### B. In-plane motion control

To control the in-plane motion of the probe we implement an image-based visual servoing where the visual features  $\mathbf{s}_2 = (u_x \ u_y \ \gamma)^T$  are directly the translation  $\mathbf{u} = (u_x \ u_y)^T$  and the rotation  $\gamma$  extracted and expressed in the current image by using the method described in section III. The corresponding desired features vector to reach is  $\mathbf{s}_2^* = (0 \ 0 \ 0)^T$  and the interaction matrix  $\mathbf{L}_{\mathbf{s}_2}$  related to  $\mathbf{s}_2$  such that  $\dot{\mathbf{s}}_2 = \mathbf{L}_{\mathbf{s}_2} \mathbf{v}_2$ , is simply a  $3 \times 3$  identity matrix. The control velocity  $\mathbf{v}_2 = (v_x \ v_y \ \omega_z)^T$  to apply to the probe in order to obtain an exponential decreasing of the visual error  $\mathbf{e}_2 = \mathbf{s}_2 - \mathbf{s}_2^*$  is then obtained by:

$$\mathbf{v}_2 = -\lambda_2 (\mathbf{L}_{\mathbf{s}_2})^{-1} \mathbf{e}_2 \text{ with gain } \lambda_2 > 0 \quad (24)$$

where  $v_x, v_y$  are the translational velocities of the probe along the  $X$  and  $Y$  axis of  $\{c\}$ , and  $\omega_z$  is the rotational velocity around the  $Z$  axes.

The 6-DOF control needed to track the full motion of the target B-scan is finally performed by applying to the probe the screw velocity  $\mathbf{v} = (v_x \ v_y \ v_z \ \omega_x \ \omega_y \ \omega_z)^T$  whose components are given by the two independent control laws (23) and (24).

## V. SIMULATION AND RESULTS

### A. Ultrasound imagery simulator

Usually, we apply simulated ground truth data to see how the system performs under ideal circumstances and then gradually introduce systemic and random errors into the data and the tracking system, thus gradually approach realistic scenarios, before an experimental validation on real data (especially on human data) is attempted. Therefore, we developed an ultrasound simulator software which allows us to position and move a 2D virtual probe and simulate a moving 3D ultrasound volume. This volume is composed from 100 parallel real B-scan images of  $180 \times 210$  pixels resolution with a pixel size of  $0.2 \times 0.2$  mm that were previously captured from an ultrasound speckle phantom at elevation intervals of 0.25 mm.

This simulator is built from the Visualization ToolKit (VTK) software system [15] and the Visual Servoing Platform (ViSP) [16] which are freely available open source C++ libraries. We use VTK to render the 3D view of the ultrasound volume as shown in Fig. 6 and to generate the current 2D ultrasound image observed by the virtual probe by means of cubic interpolation, and we use ViSP to implement the target B-scan motion extraction from the resliced 2D image and to compute the visual servoing control law applied to the probe velocity screw.

### B. Tracking robotic task results

We simulate the 6-DOF motion of the volume by applying 6 sinusoidal signals with same period of 5 seconds to the position of a Cartesian frame  $\{o\}$  attached to the volume and initially superposed to the ultrasound plane frame  $\{c\}$  such that  $\{o(t=0)\} = \{c(t=0)\}$ . The translational magnitudes were set to 10 mm along the  $X, Y$  and 12 mm along the

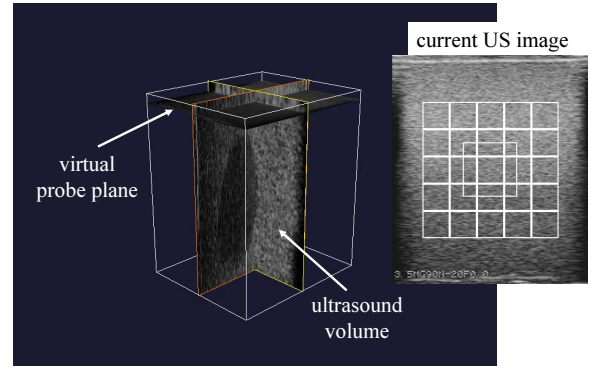


Fig. 6. Ultrasound simulator: 3D view of the ultrasound volume and the initial ultrasound image observed by the virtual probe with the 25 speckle patches (grid) and the in-plane tracking region of interest (biggest box)

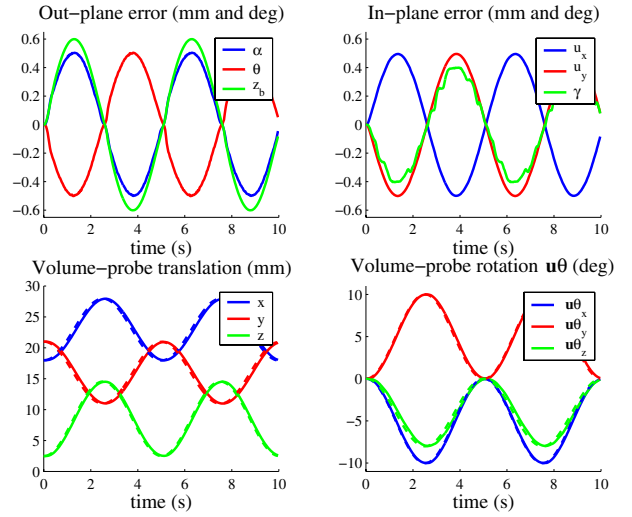


Fig. 7. (top) Out-of-plane and in-plane tracking positioning errors - (bottom) Position and orientation ( $\mathbf{u}\theta$  representation) of the volume (dashed lines) and the ultrasound probe (full lines) with respect to a fixed base frame

$Z$  axes of  $\{o\}$  and the rotational magnitudes were set to 10 deg around the  $X, Y$  axis and 8 deg around the  $Z$  axes. The threshold elevation distance  $s$  used in the state-transition graph was set to 0.1 mm and the gain of the control laws (23) and (24) were both fixed to  $\lambda_1 = \lambda_2 = 10$ .

Fig. 7 shows the time responses of the out-of-plane and in-plane positioning errors during the full motion tracking task. The components of the out-of-plane error correspond to the  $\alpha, \beta$  angles and the elevation distance  $z_b$  of the target B-scan plane with respect to the current B-scan. Their values are linked to the visual feature  $\mathbf{s}_1$  by relation (11) whereas the in-plane error corresponds directly to the visual features vector  $\mathbf{s}_2$ . Fig. 7 also shows the evolution of the volume position (dashed lines) and probe position (full lines) with respect to a fixed base frame. We can see that the task is performed well since only tracking errors lower than 0.6 mm for the translation and 0.5 deg for rotation components are measured. Moreover, these errors could be reduced if a prediction of the velocity of the target B-scan is introduced in the control law by the use for example of a Kalman filter

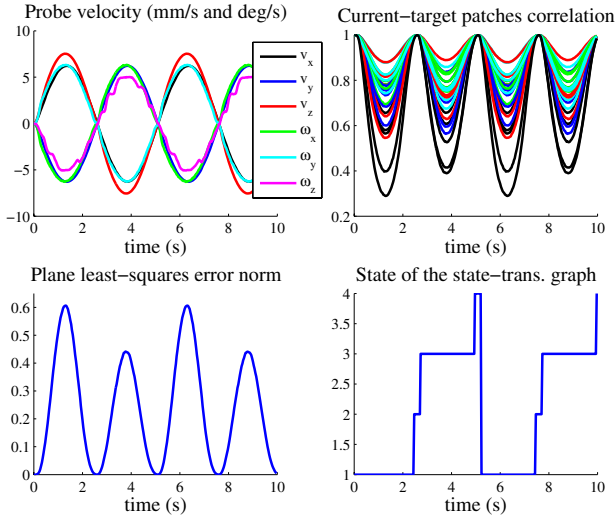


Fig. 8. (top) Velocity control screw applied to the virtual ultrasound probe and speckle correlation values of the patches between the current and target image plane - (bottom) Target plane least-squares error norm and state value of the state-transition graph use to extract the elevation sign

as presented in [17] or a generalized predictive controller [18].

Fig. 8 shows the control velocity screw applied to the probe and the evolution of the inter-patch speckle correlation values between the current and target B-scan images. The figure also presents the evolution of the plane estimation least-squares error norm and the cycle of the state-transition graph performed to track the elevation distance sign. As we can see, correlation values are decreasing due to the tracking error. Consequently if the error becomes too high, a tracking failure could occur due to a lack of correlation. The solution to avoid this case and to guarantee the robustness of the tracking robotic task is to use in the method presented in section II-C several successive intermediate planes between the current and target planes in such way that there is always a minimum speckle correlation value between themselves. This can also be useful in order to automatically retrieve an other moving B-scan target that was previously recorded at a distant position by guiding the probe along a path of intermediate B-scan images.

## VI. CONCLUSION

This paper has presented an estimation and control method to automatically synchronize the 6-DOF motion of an ultrasound probe with a moving 3D ultrasound volume by tracking the displacement of a B-scan image of reference. The out-of-plane motion was extracted from the speckle information contained in the ultrasound image, and an image region tracking method was used to extract the in-plane motion. Two independent visual control schemes were then developed to automatically move the probe in order to track the full motion of the target B-scan.

As simulations results are promising, we are now working on the experimental validation of the proposed method. The experiment will consist of controlling a medical robot

holding a 2D ultrasound probe in order to track the motion of real soft tissues. In future work, we also plan to consider non-rigid motion due to the internal deformation of soft tissues.

## ACKNOWLEDGMENTS

The authors acknowledge the support of the National Science Foundation under Engineering Research Center grant EEC9731748 and Dr. Emad Bector for his comments, discussions and providing the needed 3DUS data.

## REFERENCES

- [1] P. Abolmaesumi, S. E. Salcudean, W. H. Zhu, M. Sirouspour, S. DiMaio, Image-guided control of a robot for medical ultrasound, *IEEE Trans. Robotics and Automation*, 18:11-23, February 2002.
- [2] J. Hong, T. Dohi, M. Hashizume, K. Konishi, N. Hata, An ultrasound-driven needle insertion robot for percutaneous cholecystostomy, *Physics in Medicine and Biology*, 49(3):441-455, 2004.
- [3] J. Stoll, P. Novotny, R. Howe, P. Dupont, Real-time 3D ultrasound-based servoing of a surgical instrument, *In IEEE Int. Conf. on Robotics and Automation*, Orlando, Florida, May 2006.
- [4] M. A. Vitrani, G. Morel, T. Ortmaier, Automatic guidance of a surgical instrument with ultrasound based visual servoing, *In IEEE Int. Conf. on Robotics and Automation*, 510-515, Barcelona, Spain, April 2005.
- [5] W. Bacht, A. Krupa, Towards ultrasound image-based visual servoing, *In IEEE Int. Conf. on Robotics and Automation*, 4112-4117, Orlando, Florida, May 2006.
- [6] A. H. Gee, R. J. Housden, P. Hassenpflug, G. M. Treece, R. W. Prager, Sensorless freehand 3D ultrasound in real tissues: Speckle decorrelation without fully developed speckle, *Medical Image Analysis*, 10(2):137-149, April 2006.
- [7] R.-F. Chang, W.-J. Wu, D.-R. Chen, W.-M. Chen, W. Shu, J.-H. Lee, L.-B. Jeng, 3-D US frame positioning using speckle decorrelation and image registration, *Ultrasound in Med. & Bio.*, 29(6):801-812, 2003.
- [8] E.M Bector, I. Iordachita, G. Fichtinger, G. D. Hager, Real-time quality control of tracked ultrasound, *In 8th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*, 621-630, Palm Springs, California, USA, October 2005.
- [9] E. Bector, M. deOliveira, M. Choti, R. Ghanem, R.H. Taylor, G. Hager, G. Fichtinger, Ultrasound Monitoring of Tissue Ablation Via Deformation Model and Shape Priors, *In 9th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention*, 405-412, Denmark, 2006.
- [10] W.L Smith, A. Fenster, Optimum scan spacing for three-dimensional ultrasound by speckle statistics, *Ultrasound in Med. & Bio.*, 26(4):551-562, 2000.
- [11] G. D. Hager, P. N. Belhumeur, Efficient region tracking with parametric models of geometry and illumination, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(10):1025-1039, 1998.
- [12] S. Benhimane and E. Malis, Real-time image-based tracking of planes using efficient second-order minimization, *In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 943-948 Sendai, Japan, 2004.
- [13] B. Espiau, F. Chaumette, P. Rives, A new approach to visual servoing in robotics, *IEEE Trans. Robotics and Automation*, 8(3):313-326, June 1992.
- [14] E. Malis, Improving vision-based control using efficient second-order minimization techniques, *In IEEE Int. Conf. on Robotics and Automation*, New Orleans, USA, April 2004.
- [15] W. Schroeder, K. Martin, B. Lorensen, *The Visualization Toolkit An Object-Oriented Approach To 3D Graphics*, 3rd Edition, ISBN 1-930934-12-2 Kitware, Inc. publishers.
- [16] E. Marchand, F. Spindler, F. Chaumette, ViSP for visual servoing: a generic software platform with a wide class of robot control skills, *IEEE Robotics and Automation Magazine*, 12(4):40-52, December 2005.
- [17] F. Bensalah, F. Chaumette. Compensation of abrupt motion changes in target tracking by visual servoing, *In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, IROS'95*, 1:181-187, Pittsburgh, Pennsylvannie, August 1995.
- [18] R. Ginhoux, J. Gangloff, M. de Mathelin, L. Soler, M.M. Arenas Sanchez, J. Marescaux, Active Filtering of Physiological Motion in Robotized Surgery Using Predictive Control, *In IEEE Transactions on Robotics*, 21(1):67-79, February 2005.